

Tilburg University

Problemen rond niet-lineaire regressie

Heuts, R.M.J.; Vandaele, W.H.

Publication date:
1970

[Link to publication in Tilburg University Research Portal](#)

Citation for published version (APA):

Heuts, R. M. J., & Vandaele, W. H. (1970). *Problemen rond niet-lineaire regressie*. (EIT Research memorandum / Tilburg Institute of Economics; Vol. 8). Unknown Publisher.

General rights

Copyright and moral rights for the publications made accessible in the public portal are retained by the authors and/or other copyright owners and it is a condition of accessing publications that users recognise and abide by the legal requirements associated with these rights.

- Users may download and print one copy of any publication from the public portal for the purpose of private study or research.
- You may not further distribute the material or use it for any profit-making activity or commercial gain
- You may freely distribute the URL identifying the publication in the public portal

Take down policy

If you believe that this document breaches copyright please contact us providing details, and we will remove access to the work immediately and investigate your claim.

CBM

R

7626

1970

8

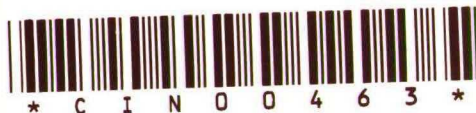
EIT

8

Bestemming	TIJDSCHRIFTENBUREAU BIBLIOTHEEK KATHOLIEKE HOGESCHOOL TILBURG	Nr.
------------	---	-----

R. M. J. Heuts en Walter H. Vandaele

**Problemen rond
niet - lineaire regressie**

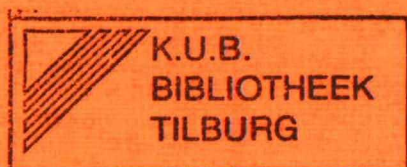


R 11

v nonlinear models

Research memorandum

ECONOMISCH INSTITUUT TILBURG
ECONOMETRISCHE AFDELING



Problemen rond niet - lineaire regressie

door

HEUTS R.M.J. en WALTER H. VANDAELE

COMP

Hout Tijdschrift Tijdschrift	S. Nr. 156.741
	Sig. R 41 (ECO)
	UDC

1. Inleiding

In de literatuur vindt men vaak een onderscheid tussen intrinsiek lineaire en intrinsiek niet-lineaire modellen (zie [7, blz. 264]). Een intrinsiek lineair model kan via een transformatie lineair in de parameters gemaakt worden, terwijl bij een intrinsiek niet-lineair model dit niet mogelijk is.

Het oplossingsalgoritme welke we in § 3 zullen behandelen geeft een oplossing voor het schatten van parameters bij intrinsiek niet-lineaire modellen.

In § 4 worden betrouwbaarheidsgebieden van de parameters van een niet-lineair model afgeleid. Hierbij wordt gebruik gemaakt van Beale's maten van niet-lineariteit. De variantie van het geschat niet-lineair model wordt in de daarna volgende paragraaf behandeld.

Van de vele economische toepassingen van niet-lineaire modellen vermelden wij er slechts enkele: de groei-modellen gebruikt voor de middellange en lange termijnvoorspelling, de prijsbepaling van apparaten in de procesindustrie (prijs niet-lineair afhankelijk van gewicht, capaciteit van de apparaten), het bepalen van een afloepkromme (learning-curve) als hulpmiddel bij controle en planning. In de algemene econometrie denken we aan het bepalen van productiefuncties: Cobb-Douglas, C.E.S.; het schatten van consumptie-, investeringsfuncties welke niet-lineair in de parameters zijn bv. door het incorporeren van verdragingspatronen. Het logistisch groeimodel wordt in § 6 nader uitgewerkt.

2. Het niet-lineaire model

De betekenis van het adjectief niet-lineair moet hier verstaan worden als niet-lineair in de parameters van een regressiemodel, niet te verwarren met niet-lineair in de verklarende variabelen x , [15, blz. 355] .

Het niet-lineaire model kan als volgt weergegeven worden:

$$\begin{aligned} y &= f(x_1, x_2, \dots, x_k; \theta_1, \theta_2, \dots, \theta_p) + \epsilon \\ &= f(x; \theta) + \epsilon \end{aligned} \quad \dots (2.1)$$

We hebben n onafhankelijke waarnemingen, zodat (2.1) geschreven wordt als:

$$y_h = f(x_h; \theta) + \epsilon_h \quad h = 1(1)n$$

Veronderstellingen:

- (i) de x_h vector is een k -dimensionale vector van onafhankelijke variabelen;
- (ii) θ is een p -dimensionale vector met componenten $\theta_1, \theta_2, \dots, \theta_p$;
- (iii) $f(x_h; \theta)$ heeft eerste afgeleiden met betrekking tot θ ;
- (iv) $\epsilon = (\epsilon_1, \epsilon_2, \dots, \epsilon_n)$ is n -dimensionaal onafhankelijk normaal verdeeld met verwachtingswaarde nul en covariantiematrix $\sigma^2 I$, $\sigma^2 > 0$ en eindig, :
 $\epsilon \sim N(0, \sigma^2 I)$.

3. Het oplossen van niet-lineaire regressieproblemen

3.1 Inleiding

De besproken methode is gebaseerd op een algoritme van Davidon, W.C. [6], aangepast door Fletcher, R. en M.J.D. Powell [8], Barnes, J.G.P.,-Rosen, E.M. [1], [18] en later gegeneraliseerd in een reeks ongepubliceerde artikelen van Shanno, D.F. [19], en Shanno, D.F. en P.C. Kettler [20]. Het is een iteratieve procedure voor het vinden van een relatief minimum van een functie van meerdere variabelen zonder bijvoorwaarden.

De standaard kwadratische vorm in p-dimensies:

$$Q(x; \theta) = Q_0 + \sum_{i=1}^p \theta_i x_i + \frac{1}{2} \sum_{i=1}^p \sum_{j=1}^p S_{ij} \theta_i \theta_j \quad \dots (3.1)$$

$$= Q_0 + x' \theta + \frac{1}{2} \theta' S \theta$$

wordt verondersteld een minimum te hebben, waarvan Fletcher, R. en M.J.D. Powell [8], blz. 164 bewijzen dat het in p-iteraties bereikt wordt. De hieronder beschreven methode bleek voor tal van problemen een zeer snelle convergentie te bezitten t.o.v. andere algorithmes.

3.2 Methode

De procedure is als volgt:

Stel we hebben een niet-lineair model (2.1) met de veronderstellingen (i), (ii) en (iii) waarvan we de kleinste kwadraatoplossing zoeken, d.w.z. die verzameling van parameters θ_j ($j = 1, 2, \dots, p$) die de criterium-functie:

$$Q(x; \theta) = \sum_{h=1}^n (y_h - f(x_h; \theta))^2 \text{ minimaliseert} \quad \dots(3.2)$$

$Q(x; \theta)$ wordt nu als volgt geminimaliseerd:

- Gegeven de vector $\theta^{(i)}$ (zie § 3.3), bereken de gradient-vector

$$g^{(i)} = \left(\frac{\partial Q(x; \theta)}{\partial \theta} \right)_{\theta=\theta^{(i)}}, \text{ een } p \times 1 \text{ matrix} \quad \dots(3.3)$$

met als j-de element:

$$-2 \sum_{h=1}^n (y_h - f(x_h; \theta_j)) f^{(i)}(x_h; \theta_j)$$

waarbij

$$f^{(i)}(x_h; \theta_j) = \left(\frac{\partial f(x_h; \theta_j)}{\partial \theta_j} \right)_{\theta_j = \theta_j^{(i)}}$$

en

i de i-de iteratie is.

- Daarna bepalen we de richting

$$s^{(i)} \equiv s^{(i)}(x; \theta) = - H^{(i)} g^{(i)} \quad \dots(3.4)$$

Op de bepaling van de matrix H zal later worden teruggekomen.

- Vervolgens zoeken we langs de lijn $s^{(i)}$ een α_i zodanig dat $Q(x; \theta^{(i)} + \alpha_i s^{(i)})$ minimaal is. Fletcher, R. en M.J.D. Powell hebben bewezen dat α_i altijd een positieve scalair mag zijn [8, blz. 164] .

3.3 Startwaarde van de te schatten parameters $\theta^{(0)}$

Voor de berekening van de startwaarden $\theta^{(0)}$ kan men als volgt te werk gaan.

Neem een willekeurig aantal parametervectoren in de p-dimensionale parameterruimte en bereken voor iedere vector in deze ruimte de som van het kwadraat van de afwijkingen: $Q(x; \theta)$, zie (3.2).

Deze vector waarvoor de som van het kwadraat van de afwijkingen minimaal is, wordt als startvector gebruikt.

3.4 Het bepalen van α_i

De bepaling van α_i geschiedt door aan de Q-waarden een polynoom van de 2-de graad in α_i aan te passen en vervolgens het minimum naar α_i van deze polynoom te bepalen. Hierbij maken we gebruik van het Gregory-Newton theorema [†] "Als er $n + 1$ equidistante punten bestaan, is het mogelijk om een polynoom van de n-de graad $P_n(x)$ of minder aan te passen, die door de $n + 1$ punten gaat".

Of in formule:

$$P_n(x) = y_0 + u[1] \Delta y_0 + u[2] \frac{\Delta^2 y_0}{2!} + \dots + u[n] \frac{\Delta^n y_0}{n!} \dots (3.5)$$

[†] WEEG, Gerard P. en George B. REED. Introduction to Numerical Analysis. Waltham, Blaisdell Publishing Company, 1966, blz. 48.

Hierbij is:

x	f(x)	$\Delta f(x)$	$\Delta^2 f(x)$	$\Delta^3 f(x)$
x_0	y_0			
		Δy_0		
x_1	y_1		$\Delta^2 y_0$	
		Δy_1		$\Delta^3 y_0$
x_2	y_2		$\Delta^2 y_1$	
		Δy_2		
x_3	y_3			

$$u^n = \frac{1}{d^n} (x-x_0)(x-x_1)\dots(x-x_{n-1}) \quad \dots (3.6)$$

d is de afstand tussen de equidistante punten.

Stel we passen aan 3 punten een 2-de graads polynoom aan:

$$P_2(x) = y_0 + u^1 \Delta y_0 + u^2 \frac{\Delta^2 y_0}{2!} \quad \dots (3.7)$$

waarbij:

$$\left. \begin{aligned} \Delta y_0 &= y_1 - y_0 \\ \Delta^2 y_0 &= y_2 - 2y_1 + y_0 \\ u^1 &= \frac{x - x_0}{d} \\ u^2 &= \frac{(x-x_0)(x-x_1)}{d^2} \end{aligned} \right\} \quad \dots (3.8)$$

ofwel:

$$P_2(x) = y_0 + \left(\frac{x-x_0}{d}\right)(y_1-y_0) + \frac{(x-x_0)(x-x_1)}{d^2} \left(\frac{y_2-2y_1+y_0}{2!}\right) \quad \dots (3.9)$$

Hiervoor wordt een extreem bereikt als:

$$x_{\min} = \frac{1}{2} (x_0 + x_1) - \frac{d(y_1 - y_0)}{(y_2 - 2y_1 + y_0)} \quad \dots (3.10)$$

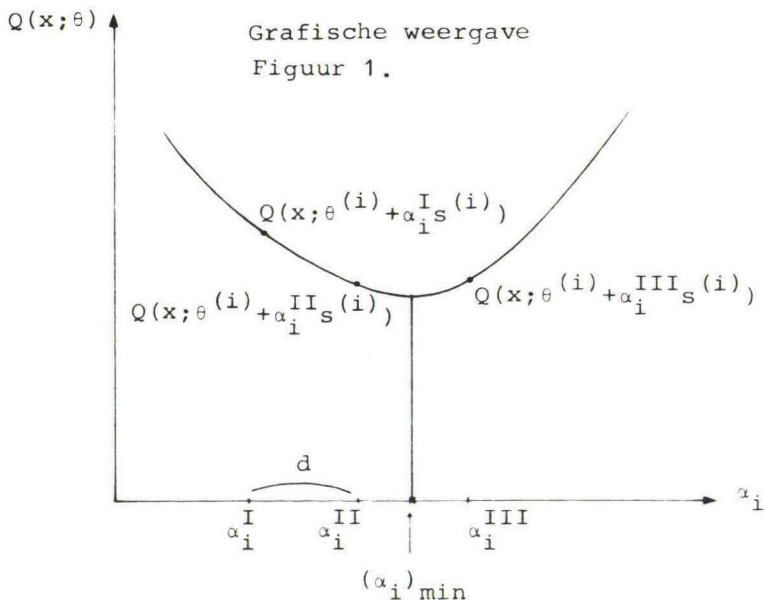
Dit extreem is een minimum als $y_2 + y_0 > 2y_1$.

Getransformeerd naar de oorspronkelijke variabelen, vinden we voor het optimum van de staplengte:

$$(\alpha_i)_{\min} = \frac{1}{2} (\alpha_i^I + \alpha_i^{II}) - \frac{d [Q(x; \theta^{(i)} + \alpha_i^{II} s^{(i)}) - Q(x; \theta^{(i)} + \alpha_i^I s^{(i)})]}{Q(x; \theta^{(i)} + \alpha_i^{III} s^{(i)}) - 2Q(x; \theta^{(i)} + \alpha_i^{II} s^{(i)}) + Q(x; \theta^{(i)} + \alpha_i^I s^{(i)})} \quad \dots (3.11)$$

Hierbij zijn α_i^I , α_i^{II} en α_i^{III} drie equidistante punten. De formule (3.11) geeft een minimum als:

$$Q(x; \theta^{(i)} + \alpha_i^{III} s^{(i)}) + Q(x; \theta^{(i)} + \alpha_i^I s^{(i)}) > 2Q(x; \theta^{(i)} + \alpha_i^{II} s^{(i)}) \quad \dots (3.12)$$



Conclusie: Men neemt drie equidistante punten α_1 waarbij de functie $Q(x; \theta)$ voldoet aan voorwaarde (3.12); past daarna een 2-de graadspolynoom naar α_1 aan en zoekt het punt waarvoor deze parabool minimaal is.

Het minimum $(\alpha_1)_{\min}$ wordt dan gebruikt om de staplengte langs de lijn $s^{(1)}$ te bepalen:

$$\sigma^{(i)} = (\alpha_1)_{\min} s^{(i)}, \quad \dots (3.13)$$

waarna de nieuwe vector van parameters kan berekend worden:

$$\theta^{(i+1)} = \theta^{(i)} + \sigma^{(i)} \quad \dots (3.14)$$

Dit algoritme wordt herhaald tot een minimum bereikt wordt (zie (3.6)).

3.5 Het bepalen van de H-matrix

De methode verschilt grotendeels van die van Hartley H.O. [13], vanwege het feit dat er noch een matrix van tweede afgeleiden, noch zijn inverse vereist is. Het is vaak voorgekomen dat na een aantal iteraties de matrix van tweede partiële afgeleiden singulier werd.

De Quasi-Newton methode[†] zal $[S]^{-1}$ door een $H^{(i)}$ matrix benaderen. De startmatrix $H^{(0)}$ mag elke positief definitieve symmetrische matrix zijn. Kiezen we als eerste startwaarde de eenheidsmatrix, dan wordt de eerste richting gekozen langs de lijn van de steilste helling. (Zie (3.4)). Uit de formules (3.4, 3.13, en 3.14) volgt:

[†] BROYDEN, C.G. [3].

$$\theta^{(i+1)} = \theta^{(i)} - \alpha_1 H^{(i)} g^{(i)}.$$

Stel nu $z^{(i)} = g^{(i+1)} - g^{(i)}$ en

$$g^{(i)} = \nabla Q = x + S \theta^{(i)} \quad \text{zie (3.1, 3.2)}$$

dan is:

$$\begin{aligned} z^{(i)} &= S(\theta^{(i+1)} - \theta^{(i)}) \\ &= S \sigma^{(i)} \end{aligned} \quad \text{zie (3.14)}$$

Voor-vermenigvuldigen met $H^{(i)}$ geeft:

$$H^{(i)} z^{(i)} = \sigma^{(i)} \quad \dots (3.15)$$

Aangezien aan de betrekking (3.15) in het algemeen niet zal voldaan zijn, wordt aangenomen dat de fout toe te schrijven is aan de berekeningswijze van $H^{(i)}$. De matrix $H^{(i)}$ bezit reeds tot op zekere hoogte de eigenschappen van de inverse van de S-matrix. Daarom wordt aan $H^{(i)}$ een correctie-matrix $D^{(i)}$ toegevoegd zodat wel aan de betrekking (3.15) is voldaan:

$$(H^{(i)} + D^{(i)}) z^{(i)} = \sigma^{(i)}$$

of:

$$D^{(i)} z^{(i)} = \sigma^{(i)} - H^{(i)} z^{(i)} \quad \dots (3.16)$$

Stel $H^{(i+1)} = H^{(i)} + D^{(i)}$, $D^{(i)}$ voldaan aan (3.16), dan hebben we:

$$H^{(i+1)} z^{(i)} = \sigma^{(i)} \quad \dots (3.17)$$

Het belang van het voldoen van de matrix $H^{(i+1)}$ aan de betrekking (3.17) ligt hierin dat:

Theorem 1 (Fletcher en Powell)

Indien Q een positief definitie kwadratische vorm is, en bij elke stap aan (3.17) is voldaan, dan zal het minimum van de functie bereikt worden in maximaal p -iteraties.

Het bewijs is gegeven in Fletcher, R. en M.J.D. Powell [8, blz. 164]. Voor algemenere functies, die niet kwadratisch zijn, maar wel bij benadering in de buurt van het minimum, geldt de convergentie van H ook, maar niet zo snel.

Om $D^{(i)}$ te berekenen kunnen we (3.16) als functie van een parameter t als volgt herschrijven:

$$D^{(i)} z^{(i)} = t \sigma^{(i)} + (1-t) \sigma^{(i)} - H^{(i)} z^{(i)} \quad \dots (3.18)$$

waaruit volgt dat:

$$D^{(i)} = t \frac{\sigma^{(i)} \sigma^{(i)'} }{\sigma^{(i)'} z^{(i)}} + \frac{[(1-t) \sigma^{(i)} - H^{(i)} z^{(i)}] [(1-t) \sigma^{(i)} - H^{(i)} z^{(i)}]'}{[(1-t) \sigma^{(i)} - H^{(i)} z^{(i)}]'} z^{(i)} \quad \dots (3.19)$$

Als $t = 0$ dan hebben we de Barnes-Rosen [1], [18] correctie en $t = 1$ is de modificatie voorgesteld door Fletcher-Powell [8].

Aldus hebben we

$$H^{(i+1)} = H^{(i)} + t \frac{\sigma^{(i)} \sigma^{(i)'}}{\sigma^{(i)' z^{(i)}}} +$$

$$\frac{(1-t) \sigma^{(i)} - H^{(i)} z^{(i)}}{(1-t) \sigma^{(i)} - H^{(i)} z^{(i)'}} \frac{(1-t) \sigma^{(i)} - H^{(i)} z^{(i)'}}{z^{(i)'}}$$

... (3.20)

Theorem 2 (Shanno)

Voor $H^{(i)}$ positief definit en $\forall t > \frac{\alpha_i - 1}{\alpha_i}$, is $H^{(i+1)}$ positief definit.

Het lang bewijs hiervan kan gevonden worden in Shanno D.F. [19, blz. 8 - 11].

Het gevolg van deze stelling is dat voor $t > \frac{\alpha_i - 1}{\alpha_i}$, $H^{(i+1)}$ positief definit is, zodat in geen enkele stap de kleinste eigenwaarde van deze matrix ooit nul kan worden. Het is evenwel mogelijk dat λ_1 de kleinste eigenwaarde van $H^{(i+1)}$ naar nul nadert:

$$\lim_{i \rightarrow \infty} \lambda_1 = 0.$$

In dit geval zal het algorithmen ontaarden voor $i \rightarrow \infty$.

Daarom zal bij elke stap, t zodanig gekozen worden dat de kleinste eigenwaarde van $H^{(i+1)}$ gemaximaliseerd wordt en tevens $t > \frac{\alpha_i - 1}{\alpha_i}$.

Dit wordt o.a. bereikt door $v' H^{(i+1)} v$, voor een willekeurige v , als functie van t te maximaliseren.

Theorem 3 (Shanno)

$v'H^{(i+1)}v$, v een willekeurige vector, is een niet - afnemende functie van t .

Uit Theorem 3 volgt dus dat de conditie van de matrix $H^{(i+1)}$ verbetert met t te laten toenemen. Daarom is het gewenst een uitdrukking voor $H^{(i+1)}$ te vinden voor $t = \infty$.

Theorem 4 (Shanno)

zij $H^{(i+1)}$ de uitdrukking (3.20)
en

$$r = \frac{\sigma^{(i)'} z^{(i)}}{\sigma^{(i)'} z^{(i)} + z^{(i)'} H^{(i)} z^{(i)}}$$

dan is

$$\begin{aligned} \lim_{t \rightarrow \infty} H^{(i+1)} &= H^{(i)} + \frac{[\sigma^{(i)} - r H^{(i)} z^{(i)}] [\sigma^{(i)} - r H^{(i)} z^{(i)}]'}{[\sigma^{(i)} - r H^{(i)} z^{(i)}]' z^{(i)}} \\ &\quad + (r-1) \frac{H^{(i)} z^{(i)} z^{(i)'} H^{(i)}}{z^{(i)'} H^{(i)} z^{(i)}} \quad \dots (3.21) \end{aligned}$$

Het algorithme voor het bepalen van het minimum van $Q(x; \theta)$ kan nu als volgt samenvattend weergegeven worden:

- Kies een startwaarde voor θ^0 en stelt $H^0 = I$

a) Bepaal $s^{(i)} = -H^{(i)}g^{(i)}$

b) Vind α_i waarvoor $Q(x; \theta^{(i)} - \alpha_i s^{(i)})$ een minimum is.

Een noodzakelijke voorwaarde hiervoor is

$$g^{(i+1)'} H^{(i)} g^{(i)} = 0$$

c) Stel $\sigma^{(i)} = \alpha_i s^{(i)}$ en vorm de vector

d) $\theta^{(i+1)} = \theta^{(i)} + \sigma^{(i)}$

e) Bepaal $H^{(i+1)}$ volgens de uitdrukking (3.20) met

$$t > \frac{\alpha_i - 1}{\alpha_i}$$

- Ga terug naar a)

3.6 Stopcriteria

Stopcriteria voor het algorithm kunnen zijn

a) Elke parameter van de vector θ verandert procentueel maar weinig meer:

$$\left| \frac{\theta_j^{(i+1)} - \theta_j^{(i)}}{\theta_j^{(i)}} \right| \leq 1.0^{-5} \quad ; \quad j = 1(1)p$$

b) De functiewaarde verandert zo goed als niet meer:

$$\left| \frac{Q(x; \theta^{(i+1)}) - Q(x; \theta^{(i)})}{Q(x; \theta^{(i)})} \right| \leq 1.0^{-5}$$

c) De afgeleide $Q^{(i)}(x; \theta)$, d.w.z. de gradient $g^{(i)}$ verschilt weinig van nul.

d) Het aantal iteraties bereikt een vooraf vastgesteld maximum.

3.7 Asymptotische covariantie-matrix van de KK-benadering

Men kan bewijzen [†] dat de asymptotische covariantiematrix van de meest aannemelijke schatters gegeven wordt door:

$$\text{cov}(\hat{\theta}) = \left[-E(l_{ij}) \right]^{-1} \quad \dots (3.23)$$

waarbij $l_{ij} = \frac{\partial^2 \ln L}{\partial \theta_i \partial \theta_j}$

en

$$\begin{aligned} \ln L &= \sum_{h=1}^n F(y_h - f(x_h; \theta)) \\ &= -\frac{n}{2} \ln(2\pi) - n \ln \sigma - \frac{1}{2\sigma^2} \sum_{h=1}^n \left[y_h - f(x_h; \theta) \right]^2 \\ &= C - \frac{Q(x; \theta)}{2\sigma^2} \end{aligned}$$

onder de veronderstelling dat de $y_h - f(x_h; \theta)$, $h = 1(1)n$, onderling onafhankelijk normaal verdeeld zijn met verwachtingswaarde nul en variantie σ^2 .

Na differentiatie van $\ln L$ naar θ vinden we voor

$$l_{ij} = -\frac{1}{2\sigma^2} \frac{\partial^2 Q(x; \theta)}{\partial \theta_i \partial \theta_j} = -\frac{1}{2\sigma^2} S_{ij}$$

(3.23) wordt dus:

$$\text{cov}(\hat{\theta}) = \left[-E(l_{ij}) \right]^{-1} = 2\sigma^2 \left[E(S_{ij}) \right]^{-1} = \text{cov}(\theta)$$

aangezien de KK-schatters dezelfde zijn als de MA-schatters (maximeren van $\ln L$ = minimeren van $Q(x; \theta)$) en dus

[†] KENDALL, M.G. en A. STUART [15, blz. 55]

$$E(\theta - \hat{\theta})(\theta - \hat{\theta})' = E(\theta - \tilde{\theta})(\theta - \tilde{\theta})'.$$

Aldus is er een verband gelegd tussen de asymptotische covariantie-matrix van de MA-schatters en de KK-schatters.

Indien we nu de verwachtingswaarde van S_{ij} , $E(S_{ij})$ benaderen door de waargenomen waarde S_{ij} , dan wordt

$$\text{cov}(\hat{\theta}) \approx 2\sigma^2 [S_{ij}]^{-1} \quad \dots (3.24)$$

Een schatter voor σ^2 is:

$$\hat{\sigma}^2 = \frac{Q(x; \hat{\theta})}{n-k}.$$

Sommige auteurs delen door n , i.p.v. $n-k$. De matrix $[S_{ij}]^{-1}$ is gevonden via de matrix $H^{(i+1)}$ van de laatste iteratie, zie (3.20).

4. Het betrouwbaarheidsgebied bij niet-lineaire modellen

In § 3 werd het volgende probleem aangepakt: Het vinden van KK-schatters voor een niet-lineair model (2.1), nl. door het minimaliseren van de criteriumfunctie (3.2), de som van de kwadraten van de resttermen:

$$Q(x; \theta) = \sum_{h=1}^n (y_h - f(x_h; \theta))^2 \quad \text{zie (3.2)}$$

In deze paragraaf gaan we uit van een KK-schatter $\hat{\theta}$ van θ en stellen ons de vraag hoe een theoretisch verantwoord betrouwbaarheidsgebied voor $\hat{\theta}$ gevonden kan worden.

Hierbij zal worden uitgegaan van de formule voor het vinden van betrouwbaarheidsgebieden bij een lineair model. Beale, E.M. [2] heeft immers kunnen aantonen dat het betrouwbaarheidsgebied bij een niet-lineair model gevormd wordt door het lineair betrouwbaarheidsgebied te corrigeren met een factor welke van de mate van niet-lineariteit afhangt. Daarom beginnen we deze paragraaf met een uitleg over de maten van niet-lineariteit.

Er wordt vanuit gegaan dat men vertrouwd is met de begrippen: waarnemingsruimte, schattingsruimte (oplossingsruimte), tangensruimte aan de schattingsruimte in een punt θ_0 , parameter ruimte. Verwezen kan worden naar Malinvaud, E. [16, Hoofdstukken 5 en 9]. Ook het boek van Draper, N.R. en H. Smith bevat een goede uitleg voor het gebruik ervan in niet-lineaire modellen [7, blz. 295 - 299].

4.1 Beale's maten voor niet-lineariteit

Maten voor niet-lineariteit zijn ontwikkeld door Beale, E.M.L. [2] (zie ook Guttman, I. en D.A. Meeter [11]) om te onderzoeken of de graad van niet-lineariteit in een niet-lineair regressieprobleem klein genoeg is om de resultaten van het lineaire model als benadering te mogen gebruiken.

Deze zijn:

A. Empirische maat voor niet-lineariteit

$$\hat{N}_{\theta} = \frac{ps^2 \sum_{w=1}^m || f(x; \theta_w) - t(\theta_w) ||^2}{\sum_{w=1}^m || f(x; \theta_w) - f(x; \hat{\theta}) ||^4} \quad \dots (4.1)$$

B. Theoretische maat voor niet-lineariteit

$$N_{\theta} = \frac{p\sigma^2 \sum_{w=1}^{\infty} || f(x; \theta_w) - t(\theta_w) ||^2}{\sum_{w=1}^{\infty} || f(x; \theta_w) - f(x; \hat{\theta}) ||^4} \quad \dots (4.2)$$

C. Empirische minimale maat voor niet-lineariteit

$$\hat{N}_{\phi} = \frac{ps^2 \sum_{w=1}^m || f(x; \theta_w) - t_{\phi}(\theta_w) ||^2}{\sum_{w=1}^m || f(x; \theta_w) - f(x; \hat{\theta}) ||^4} \quad \dots (4.3)$$

D. Theoretische minimale maat voor niet-lineariteit

$$N_{\phi} = \frac{p\sigma^2 \sum_{w=1}^{\infty} || f(x; \theta_w) - t_{\phi}(\theta_w) ||^2}{\sum_{w=1}^{\infty} || f(x; \theta_w) - f(x; \hat{\theta}) ||^4} \quad \dots (4.4)$$

We gaan uit van het model (2.1)

$$y = f(x; \theta) + \varepsilon \quad \text{zie (2.1)}$$

met de veronderstellingen (i), (ii), (iii) en (iv). Hier wordt dus aangenomen dat $\varepsilon \sim N(0, \sigma^2 I)$.

Als $\hat{\theta}$ de vector van de kleinste kwadraten schatters is, dan wordt een gelineariseerde schattingsruimte in de vorm van een **tangensruimte** in het punt $\hat{\theta}$ als volgt gevonden:

$$t_h(\theta) = f(x_h; \hat{\theta}) + \sum_{j=1}^p (\theta_j - \hat{\theta}_j) \left. \frac{\partial f(x_h; \theta)}{\partial \theta_j} \right|_{\theta = \hat{\theta}}$$

$$h = 1(1)n \quad \dots (4.5)$$

of

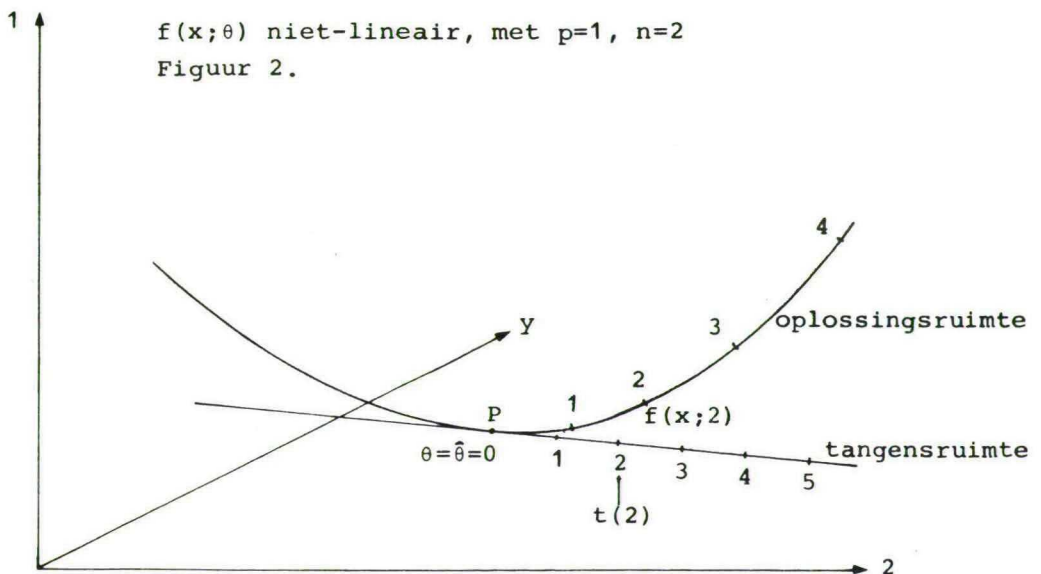
$$t(\theta) = f(x; \hat{\theta}) + X(\theta - \hat{\theta}) \quad \dots (4.6)$$

met

$$X = \left. \frac{\partial f(x; \theta)}{\partial \theta} \right|_{\theta = \hat{\theta}} = \left\{ \left. \frac{\partial f(x_h; \theta)}{\partial \theta_j} \right|_{\theta = \hat{\theta}} \right\}, \text{ een } n \times p$$

matrix (zie ook (3.3)).

Dit begrip kan geometrisch als volgt verduidelijkt worden. De waarnemingsruimte voor een niet-lineaire functie met $n = 2$ en $p = 1$ is weergegeven in figuur 2. De oplossingsruimte is de kromme met de punten $\{ f(x_1; \theta), f(x_2; \theta) \}$ als functie van θ , waarbij x_1, x_2 vast zijn. P is het punt in de oplossingruimte gelegen op de kleinste afstand van y



In figuur (2) is aangenomen dat $\hat{\theta} = 0$ is en het punt 1 op de oplossingsruimte is het punt bereikt als $\theta = 1$ is in de niet-lineaire functie. Het feit dat 1, 2, 3, ... op de oplossingsruimte geen equidistante punten zijn is te wijten aan het niet-lineair zijn van het model.

De punten op de tangensruimte zijn wel equidistant aangezien dit de eenheidstoename voorstelt van de functie in $\hat{\theta}$.

Veronderstel we bekijken de punten θ_w , $w = 1(1)m$, (waarbij elke θ_w een p-dimensionale vector is), in de buurt van de kleinste kwadraatoplossing $\hat{\theta}$, dan is volgens Beale, E.M.L. een ruwe maat voor niet-lineariteit:

$$Q_{\theta} = \sum_{w=1}^m \sum_{h=1}^n \left[f(x_h, \theta_w) - f(x_h; \hat{\theta}) - \sum_{j=1}^p \{ (\theta_w)_j - \hat{\theta}_j \} \frac{\partial f(x_h; \theta)}{\partial \theta_j} \bigg|_{\theta = \hat{\theta}} \right]^2$$

$$= \sum_{w=1}^m || f(x; \theta_w) - t(\theta_w) ||^2 \quad \dots (4.7)$$

Q_{θ} is dus de som van het kwadraat van de afstanden van de punten $f(x; \theta_w)$ op de oplossingsruimte t.o.v. de punten $t(\theta_w)$ op de tangensruimte in $f(x; \hat{\theta})$. In de praktijk kan men als waarde voor θ_w gebruiken de waarden die reeds gevonden zijn tijdens de iteratieve procedure voor het vinden van $\hat{\theta}$.

Het is duidelijk dat deze maat Q_{θ} afhankelijk is van het aantal punten $f(x; \theta_w)$ dat men gebruikt en hun afstanden t.o.v. $f(x; \hat{\theta})$. Daarom stelt Beale, E.M.L. voor Q_{θ} te normaliseren tot de empirische maat van niet-lineariteit \hat{N}_{θ}

$$\hat{N}_\theta = \frac{\sum_{w=1}^m || f(x; \theta_w) - t(\theta_w) ||^2}{\sum_{w=1}^m || f(x; \theta_w) - f(x; \hat{\theta}) ||^4} \quad \text{zie (4.1)}$$

Hierin stellen we

$$\begin{aligned} D &\equiv \sum_{w=1}^m || f(x; \theta_w) - f(x; \hat{\theta}) ||^4 \\ &= \sum_{w=1}^m \sum_{h=1}^n \{ f(x_h; \theta_w) - f(x_h; \hat{\theta}) \}^2 \quad \dots (4.8) \end{aligned}$$

en

s^2 is de beste schatter voor σ^2

Beale, E.M.L. definieert \hat{N}_θ , de genormaliseerde maat voor niet-lineariteit, als de empirische maat voor niet-lineariteit. \hat{N}_θ heeft de volgende eigenschappen:

1. Het is een dimensieloze grootheid: daar Q_θ en s^2 elk van de dimensie twee in de waarnemingen zijn en D de dimensie vier heeft;
2. Deze maat wordt slechts in zeer geringe mate beïnvloed door w , het aantal punten' gekozen rond $\hat{\theta}$;
3. De afstanden van $f(x; \hat{\theta})$ tot $f(x; \theta_w)$ in de waarnemingsruimte hebben geen invloed meer, indien deze afstanden niet al te groot zijn.

Analoog met de empirische maat voor niet-lineariteit \hat{N}_θ is er een theoretische maat voor niet-lineariteit N_θ , zie (4.2). Deze wordt verkregen door in \hat{N}_θ , s^2 te vervangen door σ^2 en het aantal waarden θ_w naar oneindig te laten toenemen.

De empirische maat voor niet-lineariteit \hat{N}_θ kan nu als volgt aangewend worden om te toetsen of een niet-line-

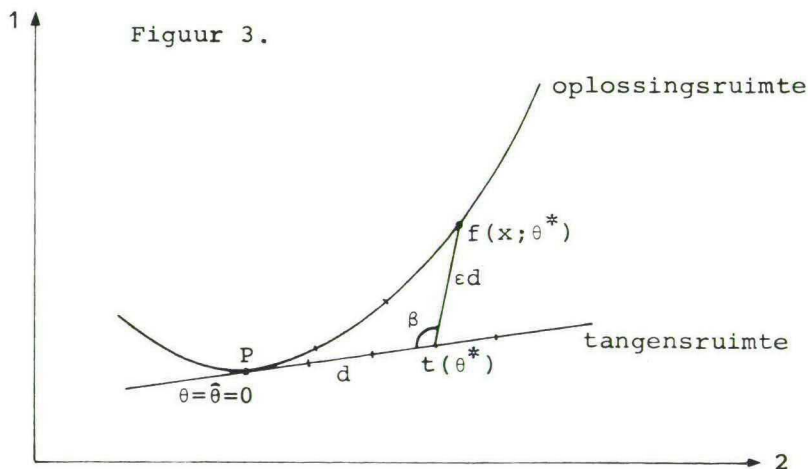
air model in de buurt van het optimum voldoende door een lineair model te benaderen is.

Bij een lineair model wordt het $100(1-\alpha)\%$ betrouwbaarheidsgebied in de waarnemingsruimte weergegeven door de p -dimensionale ellipsoïde met middelpunt P gedefinieerd als:

$$E = \{ \theta \mid S(\theta) - Q(x; \hat{\theta}) \leq ps^2 F_{\alpha}(p, n-p) \} \dots (4.9)$$

waar

s^2 een schatter van σ^2 met $n-p$ vrijheidsgraden, $F_{\alpha}(p, n-p)$ is het $100(1-\alpha)\%$ punt van een F verdeling met p , (het aantal te schatten parameters), en $n-p$ vrijheidsgraden, en $S(\theta)$ de afstand tussen y en $t(\theta)$. Aangezien $Q(x; \hat{\theta})$ ook in de tangensruimte ligt is $Q(x; \hat{\theta}) = S(\hat{\theta})$



De rand van deze verzameling vinden we door θ^* zodanig te kiezen dat $t(\theta^*)$ op een afstand

$$d = (ps^2 F_{\alpha}(p, n-p))^{\frac{1}{2}} \quad \dots (4.10)$$

van $f(x; \hat{\theta})$ ligt in de waarnemingsruimte.

Indien het model (2.1) volledig lineair zou zijn, dan was de afstand ϵd in figuur 3 nul.

We weten dat

$$||f(x; \theta_w) - f(x; \hat{\theta})||^2 = d^2 + (\epsilon d)^2 - 2 \epsilon d^2 \cos \beta$$

Voor voldoende kleine ϵ , kan elke term van $D(4.8)$ benaderd worden door d^4 ,
en (4.1) wordt

$$\hat{N}_{\theta} \approx \frac{ps^2 (\epsilon d)^2}{d^4} = \frac{ps^2 \epsilon^2}{d^2}$$

zodat

$$\epsilon^2 \approx \frac{\hat{N}_{\theta} d^2}{ps^2} = \hat{N}_{\theta} F_{\alpha}(p, n-p)$$

Beale, E.M.L. neemt nu aan dat de lineaire benadering toegelaten is als $\epsilon^2 < 0.01$, d.w.z. $t(\theta^*)$ ligt op een afstand van $f(x; \theta^*)$ welke kleiner is dan een tiende van de afstand $f(x, \theta^*) - f(x, \hat{\theta})$. Aldus leidt men af dat als

$$\hat{N}_{\theta} < 0.01 / F_{\alpha}(p, n-p) \quad \dots (4.11)$$

een lineaire benadering toegelaten is. Is evenwel

$$\hat{N}_{\theta} > 1/F_{\alpha}(p, n-p) \quad \dots (4.12)$$

dan verwerpt Beale de linearisatie.

Indien $\frac{0,01}{F_{\alpha}(p, n-p)} < \hat{N}_{\theta} < \frac{1}{F_{\alpha}(p, n-p)}$ dan moet gebruik

gemaakt worden van de empirische minimale maat voor niet-lineariteit \hat{N}_{ϕ} (4.3).

Hierop wordt verder teruggekomen.

Beale, E.M.L. heeft verder bewezen dat er een theoretische minimale maat voor niet-lineariteit (4.4) bestaat corresponderende met de theoretische maat voor niet-lineariteit gedefinieerd in (4.1).

Deze theoretische maat verkrijgt men via een parametertransformatie $\phi = \phi(\theta)$ zodat

$$\begin{aligned} t_{\phi}(\theta) &= f(x_h; \hat{\theta}) + \sum_{j=1}^p \phi_j \left. \frac{\partial f(x_h; \theta)}{\partial \theta_j} \right|_{\theta = \hat{\theta}}; h=1(1)n \\ &= f(x; \hat{\theta}) + X \phi \quad \text{zie (4.6), ... (4.13)} \end{aligned}$$

Deze parametertransformatie houdt in dat we uit de oplossingsruimte loodrecht projecteren op de tangensruimte. Zodoende krijgen we:

$$N_{\phi} = \frac{p \sigma^2 \sum_{w=1}^{\infty} || f(x; \theta_w) - t_{\phi}(\theta_w) ||^2}{\sum_{w=1}^{\infty} || f(x; \theta_w) - f(x; \hat{\theta}) ||^4} \quad \text{zie (4.4)}$$

het minimum van N_{θ} .

Deze N_{ϕ} is dus de minimale waarde welke we voor N_{θ} kunnen krijgen indien het model (2.1) als gegeven wordt beschouwd, maar een parametertransformatie toegelaten is.

De empirische minimale maat voor niet-lineariteit \hat{N}_ϕ wordt als volgt berekend. Het analogon van (4.7) is:

$$Q_\phi = \sum_{w=1}^m ||f(x; \theta_w) - t_\phi(\theta_w)||^2 \quad \dots(4.14)$$

De transformatie $\phi = \phi(\theta)$ wordt zodanig gekozen dat Q_ϕ minimaal wordt m.a.w. door het oplossen van het stelsel van normaalvergelijkingen voor elke w :

$$X' [f(x; \theta_w) - f(x; \hat{\theta}) - X\phi_w] = 0 \quad w = 1(1)m$$

Uit deze p lineaire vergelijkingen volgen de ϕ_w voor elke w . Aldus is \hat{N}_ϕ bepaald

$$\hat{N}_\phi = ps^2 Q_\phi / \sum_{w=1}^m ||f(x; \theta_w) - f(x; \hat{\theta})||^4 \text{ zie (4.3)}$$

Aangezien \hat{N}_ϕ tevens de minimale waarde is van \hat{N}_θ , kan nu, voor het geval men met \hat{N}_θ tot geen uitspraak kan komen, \hat{N}_ϕ aan dezelfde toetsen onderworpen worden. Voldoet \hat{N}_ϕ aan (4.11), dan betekent dit dat de niet-lineariteit via een parametertransformatie binnen aanvaardbare proporties teruggebracht kan worden. Is evenwel $\hat{N}_\phi > 0.01/F_\alpha(p, n-p)$ dan kan men geenszins de lineaire methoden zinvol op de niet-lineaire functie toepassen.

4.2 Het betrouwbaarheidsgebied bij niet-lineaire modellen

Een benadering van het betrouwbaarheidsgebied in het niet-lineaire model, gelegen in de n -dimensionale waarnemingsruimte wordt weergegeven door

$$E_b = \{ \theta \mid Q(x; \theta) - Q(x; \hat{\theta}) \leq ps^2 F_\alpha(p, n-p) \} \quad \dots(4.15)$$

$$Q(x; \theta) = \sum_{h=1}^n \left[y_h - f(x_h; \theta) \right]^2, \text{ zie (3.1)}$$

Beale, E.M.L. heeft bewezen 2, Appendix 1 dat voor N_ϕ niet te groot ($N_\phi < 0.01/F_\alpha(p, n-p)$) de rechterzijde van de ongelijkheid in de verzameling (4.15) vermenigvuldigd dient te worden met:

$$\left. \begin{array}{ll} 1 + N_\phi & (p = 1) \\ 1 + \frac{p+2}{p} N_\phi & (p > 2) \end{array} \right\} \dots (4.16)$$

om een betrouwbaarheidsgebied groter dan of gelijk aan $100(1-\alpha)\%$ te krijgen. In de toepassingen wordt \hat{N}_ϕ als schatter van N_ϕ gebruikt.

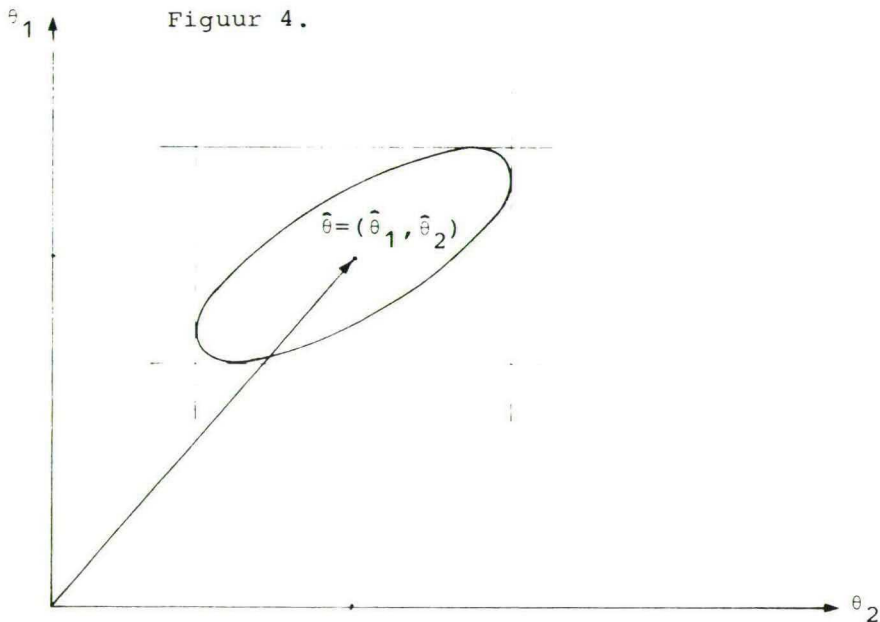
Als het aantal parameters groot is, heeft het bepalen van simultane betrouwbaarheidsgebieden volgens (4.15), (4.16) slechts een beperkte bruikbaarheid.

Teneinde een beter inzicht te krijgen in het simultaan betrouwbaarheidsgebied, gaan we eerst dit gebied uit de n-dimensionale waarnemingsruimte afbeelden in de p-dimensionale parameterruimte. Bij het lineaire model wordt het betrouwbaarheidsgebied in de parameter-ruimte weergegeven door:

$$\begin{aligned} E_p &= \{ \theta \mid (\theta - \hat{\theta})' X'X(\theta - \hat{\theta}) \leq ps^2 F_\alpha(p, n-p) \} \\ &= \{ \theta \mid (\theta - \hat{\theta})' X'X(\theta - \hat{\theta}) \leq \frac{p}{n-p} Q(x; \hat{\theta}) F_\alpha(p, n-p) \} \\ &\dots (4.17) \end{aligned}$$

hetgeen een ellipsoïde is met middelpunt $\hat{\theta}$ (Zie Goldberger, A.S. [10, blz. 178], Draper, N.R. en H. Smith, [7, blz. 286])

Met behulp van de F projectie (Scheffé) kan men deze ellipsoïde insluiten door $2p$ hypervlakken (lineaire variëteiten), loodrecht op de coördinaatassen. Voor een ellips in de 2-dimensionale parameter ruimte kan men dit als volgt illustreren:



Het is duidelijk dat als een punt binnen de ellipsoïde ligt, het zeker binnen het rechthoekig gebied ligt, maar het omgekeerde is niet waar. Daarom is het betrouwbaarheidsgebied van het rechthoekig gebied groter dan $100(1-\alpha)\%$.

In de Appendix is aangetoond dat de afwijking bij elke parameter t.o.v. zijn kleinste kwadraat gelijk is aan:

$$\theta_j - \hat{\theta}_j = \pm \sqrt{ps^2 F_\alpha(p, n-p) (X'X)^{jj}}$$

$$= \pm \sqrt{\frac{p}{n-p} Q(x; \hat{\theta}) F_\alpha(p, n-p) (X'X)^{jj}}$$

$$j = 1(1)p \dots (4.18)$$

waarbij $(X'X)^{jj}$ het j-de diagonaal element is van de matrix $(X'X)^{-1}$, en θ_j een punt op de rand van de p-dimensionale ellipsoïde is, zodat $|\theta_j - \hat{\theta}_j|$ maximaal is.

Rekening houdend met de door Beale, E.M.L. geïntroduceerde correctiefactor luidt het aangepaste betrouwbaarheidsgebied bij het niet-lineair model:

$$P \left\{ \begin{array}{l} \hat{\theta}_1 - \sqrt{k(X'X)^{11}} \leq \theta_1 \leq \hat{\theta}_1 + \sqrt{k(X'X)^{11}} \\ \hat{\theta}_2 - \sqrt{k(X'X)^{22}} \leq \theta_2 \leq \hat{\theta}_2 + \sqrt{k(X'X)^{22}} \\ \vdots \\ \hat{\theta}_p - \sqrt{k(X'X)^{pp}} \leq \theta_p \leq \hat{\theta}_p + \sqrt{k(X'X)^{pp}} \end{array} \right\} \geq 100(1-\alpha)\% \quad \dots (4.19)$$

waarbij

$$k = \frac{p}{n-p} Q(x; \hat{\theta}) F_{\alpha}(p, n-p) \beta$$

en

β de correctie factor (4.16).

5 De variantie van het geschat niet-lineair model

Stel we gaan uit van het niet-lineair model

$$y = f(x; \theta) + \varepsilon \quad \dots(2.1)$$

met veronderstellingen (i, ii, iii, iv)

De geschatte waarde voor y is:

$$\hat{y} = f(x; \hat{\theta})$$

Indien de linearisatie van het niet-lineair model verantwoord is, dan wordt een goede benadering van de variantie rond de voorspelde waarde gegeven door [†]

$$\text{VAR}(\hat{y}_h) \approx \sum_{j=1}^p \left(\frac{\partial f(x_h; \hat{\theta})}{\partial \hat{\theta}_j} \right)^2 \text{var}(\hat{\theta}_j) + 2 \sum_{j=1}^p \sum_{r=1}^p \left(\frac{\partial f(x_h; \hat{\theta})}{\partial \hat{\theta}_j} \right)$$

$$\left(\frac{\partial f(x_h; \hat{\theta})}{\partial \hat{\theta}_r} \right) \text{cov.} (\hat{\theta}_j, \hat{\theta}_r) \quad \dots(5.1)$$

[†] SVESHNIKOV, A.A. [21, blz. 136], HEUTS, R. [14, blz. 16].

6. Toepassing: Het logistische groeimodel met additieve storingsterm

6.1 Model:

$$y_h = \frac{\kappa}{1 + \beta e^{-\alpha h}} + \varepsilon_h \quad h = 1(1)n \quad \dots(6.1)$$

met $\varepsilon \sim N(0, \sigma^2 I)$ zie veronderstelling (2.1, iv)

De aannemelijkheidsfunctie is

$$L = \frac{1}{(2\pi\sigma^2)^{n/2}} \exp \left\{ -\frac{1}{2\sigma^2} \sum_{h=1}^n \varepsilon_h^2 \right\} \quad \dots(6.2)$$

$$= \frac{1}{(2\pi\sigma^2)^{n/2}} \exp \left\{ -\frac{1}{2\sigma^2} \sum_{h=1}^n \left(y_h - \frac{\kappa}{1 + \beta e^{-\alpha h}} \right)^2 \right\} \quad \dots(6.3)$$

De Jacobiaan van de transformatie (6.2) naar (6.3) is 1.

$$\ln L = -\frac{n}{2} \ln 2\pi - \frac{n}{2} \ln \sigma^2 - \frac{1}{2\sigma^2} \sum_{h=1}^n \left(y_h - \frac{\kappa}{1 + \beta e^{-\alpha h}} \right)^2 \quad \dots(6.4)$$

6.2 De meest aannemelijke schatters van κ , β , α worden gevonden door de partiële afgeleiden naar κ , β , α nul te stellen:

$$\left. \begin{aligned} \frac{\partial \ln L}{\partial \kappa} &= \frac{1}{\sigma^2} \sum_{h=1}^n \left\{ \frac{1}{1+\beta e^{-\alpha h}} \left(y_h - \frac{\kappa}{1+\beta e^{-\alpha h}} \right) \right\} \\ \frac{\partial \ln L}{\partial \beta} &= -\frac{1}{\sigma^2} \sum_{h=1}^n \left\{ \frac{\kappa e^{-\alpha h}}{(1+\beta e^{-\alpha h})^2} \left(y_h - \frac{\kappa}{1+\beta e^{-\alpha h}} \right) \right\} \\ \frac{\partial \ln L}{\partial \alpha} &= \frac{1}{\sigma^2} \sum_{h=1}^n \left\{ \frac{\kappa \beta h e^{-\alpha h}}{(1+\beta e^{-\alpha h})^2} \left(y_h - \frac{\kappa}{1+\beta e^{-\alpha h}} \right) \right\} \end{aligned} \right\} \dots (6.5)$$

Het is duidelijk dat het maximeren van $\ln L$ naar κ , β en α gelijk staat met het minimeren van de kleinste kwadraatsom

$$Q(x; \theta) = \sum_{h=1}^n \left(y_h - \frac{\kappa}{1+\beta e^{-\alpha h}} \right)^2 \quad \dots (6.6)$$

De kleinste kwadraatschatters $\hat{\kappa}$, $\hat{\beta}$, $\hat{\alpha}$ zijn bij de gestelde veronderstellingen dus gelijk aan de meest aannemelijke schatters $\tilde{\kappa}$, $\tilde{\beta}$ en $\tilde{\alpha}$. Voor de oplossing van het vergelijkingstelsel (6.5) wordt verwezen naar § 3.

6.3 De asymptotische covariantie-matrix van de K.K. schatter is:

$$\text{cov}(\hat{\theta}) = 2\sigma^2 \left[E(S_{ij}) \right]^{-1} \approx 2\hat{\sigma}^2 (S_{ij})^{-1} \quad \text{zie (3.24)}$$

De $(S_{ij})^{-1}$ van deze covariantie-matrix kan met de procedure uit § 3. banaderd worden, (de matrix $H^{(i+1)}$ van de laatste iteratie), maar de verwachting van de S_{ij} kan ook als volgt expliciet opgeschreven worden:

$$\{ E(S_{ij}) \}^{-1} =$$

$$\begin{bmatrix} 2 \sum_{h=1}^n \left(\frac{1}{1+\beta e^{-\alpha h}} \right)^2 & -2 \sum_{h=1}^n \frac{\kappa e^{-\alpha h}}{(1+\beta e^{-\alpha h})^3} & 2 \sum_{h=1}^n \frac{\kappa \beta h e^{-\alpha h}}{(1+\beta e^{-\alpha h})^3} \\ -2 \sum_{h=1}^n \frac{\kappa e^{-\alpha h}}{(1+\beta e^{-\alpha h})^3} & 2 \sum_{h=1}^n \frac{\kappa^2 e^{-2\alpha h}}{(1+\beta e^{-\alpha h})^4} & -2 \sum_{h=1}^n \frac{\kappa^2 \beta h e^{-2\alpha h}}{(1+\beta e^{-\alpha h})^4} \\ 2 \sum_{h=1}^n \frac{\kappa \beta h e^{-\alpha h}}{(1+\beta e^{-\alpha h})^3} & -2 \sum_{h=1}^n \frac{\kappa^2 \beta h e^{-2\alpha h}}{(1+\beta e^{-\alpha h})^4} & 2 \sum_{h=1}^n \frac{\kappa^2 \beta^2 h^2 e^{-2\alpha h}}{(1+\beta e^{-\alpha h})^4} \end{bmatrix}^{-1}$$

... (6.7)

De matrix (6.7) voorvermenigvuldigd met $2\sigma^2$ geeft bij benadering $\text{cov}(\hat{\theta})$.

6.4 Vervolgens gaan we onderzoeken of linearisatie van het model verantwoord is.

Een lineaire benadering van iedere theoretische verwachte waarde in de buurt van $\hat{\theta} = (\hat{\kappa}, \hat{\beta}, \hat{\alpha})$ is te schrijven als:

$$t_h(\theta) = \frac{\hat{\kappa}}{1+\hat{\beta}e^{-\hat{\alpha}h}} + (\kappa-\hat{\kappa}) \left. \frac{\partial f(x;\theta)}{\partial \kappa} \right|_{\theta=\hat{\theta}} + (\beta-\hat{\beta}) \left. \frac{\partial f(x;\theta)}{\partial \beta} \right|_{\theta=\hat{\theta}} + (\alpha-\hat{\alpha}) \left. \frac{\partial f(x;\theta)}{\partial \alpha} \right|_{\theta=\hat{\theta}}$$

... (6.8)

Neem nu waarden voor κ , β en α in de buurt van $\hat{\kappa}$, $\hat{\beta}$ en $\hat{\alpha}$ nl κ_w , β_w en α_w , $w = 1(1)m$, dan is een ruwe maat voor niet-lineariteit:

$$Q_{\kappa, \beta, \alpha} = \sum_{w=1}^m \sum_{h=1}^n \left[\frac{\kappa_w}{1+\beta_w e^{-\alpha_w h}} - \frac{\hat{\kappa}}{1+\hat{\beta} e^{-\hat{\alpha} h}} - \frac{1}{1+\hat{\beta} e^{-\hat{\alpha} h}} (\kappa_w - \hat{\kappa}) \right. \\ \left. + \frac{\hat{\kappa} e^{-\hat{\alpha} h}}{(1+\hat{\beta} e^{-\hat{\alpha} h})^2} (\beta_w - \hat{\beta}) - \frac{\hat{\kappa} \hat{\beta} h e^{-\hat{\alpha} h}}{(1+\hat{\beta} e^{-\hat{\alpha} h})^2} (\alpha_w - \alpha) \right]^2 \quad \dots (6.9)$$

De genormaliseerde maat voor niet-lineariteit is:

$$\hat{N}_{\kappa, \beta, \alpha} = \frac{3s^2 Q_{\kappa, \beta, \alpha}}{\sum_{w=1}^m \left\{ \sum_{h=1}^n \left[\frac{\kappa_w}{1+\beta_w e^{-\alpha_w h}} - \frac{\hat{\kappa}}{1+\hat{\beta} e^{-\hat{\alpha} h}} \right]^2 \right\}} \quad \dots (6.10)$$

Als aan voorwaarde (4.11) voldaan is, of $\hat{N}_{\kappa, \beta, \alpha}$ tussen de grenzen (4.11) en (4.12) ligt gaan we de empirische minimale maat voor niet-lineariteit berekenen. Zoals eerder uitgelegd lossen we daartoe het volgende stelsel van normaal vergelijkingen op voor elke w

$$X' \begin{bmatrix} \frac{\kappa_w}{1+\beta_w e^{-\alpha_w}} - \frac{\hat{\kappa}}{1+\hat{\beta} e^{-\hat{\alpha}}} \\ \frac{\kappa_w}{1+\beta_w e^{-2\alpha_w}} - \frac{\hat{\kappa}}{1+\hat{\beta} e^{-2\hat{\alpha}}} \\ \vdots \\ \frac{\kappa_w}{1+\beta_w e^{-n\alpha_w}} - \frac{\hat{\kappa}}{1+\hat{\beta} e^{-n\hat{\alpha}}} \end{bmatrix} = X' X \begin{bmatrix} \phi(\kappa) \\ \phi(\beta) \\ \phi(\alpha) \end{bmatrix}$$

waarbij

X een (nx3) matrix:

$$X = \begin{bmatrix} \frac{1}{1+\hat{\beta}e^{-\hat{\alpha}}} & \frac{-\hat{\kappa}e^{-\hat{\alpha}}}{(1+\hat{\beta}e^{-\hat{\alpha}})^2} & \frac{\hat{\kappa}\hat{\beta}e^{-\hat{\alpha}}}{(1+\hat{\beta}e^{-\hat{\alpha}})^2} \\ \frac{1}{1+\hat{\beta}e^{-2\hat{\alpha}}} & \frac{-\hat{\kappa}e^{-2\hat{\alpha}}}{(1+\hat{\beta}e^{-2\hat{\alpha}})^2} & \frac{2\hat{\kappa}\hat{\beta}e^{-2\hat{\alpha}}}{(1+\hat{\beta}e^{-2\hat{\alpha}})^2} \\ \vdots & \vdots & \vdots \\ \frac{1}{1+\hat{\beta}e^{-n\hat{\alpha}}} & \frac{-\hat{\kappa}e^{-n\hat{\alpha}}}{(1+\hat{\beta}e^{-n\hat{\alpha}})^2} & \frac{\hat{\kappa}\hat{\beta}ne^{-n\hat{\alpha}}}{(1+\hat{\beta}e^{-n\hat{\alpha}})^2} \end{bmatrix} \quad \dots (6.11)$$

De gevonden waarden voor de Φ -vector, worden in Q_Φ (4.3) gesubstitueerd om vervolgens \hat{N}_Φ te berekenen

$$\hat{N}_\Phi = \frac{3s^2 Q_\Phi}{\sum_{w=1}^m \left\{ \sum_{h=1}^n \left[\frac{\kappa_w}{1+\beta_w e^{-\alpha_w h}} - \frac{\hat{\kappa}}{1+\hat{\beta}e^{-\hat{\alpha}h}} \right]^2 \right\}^2} \quad \dots (6.12)$$

Indien (6.10) geen uitspraak geeft m.b.t. de lineaire benadering, wordt (6.12) getoetst volgens het criterium (4.11).

Als de linearisatie verantwoord is, kan als benaderend $100(1-\alpha)\%$ betrouwbaarheidsgebied voor θ de volgende drie dimensionale ellipsoïde gebruikt worden

$$\sum_{h=1}^n \left(y_h - \frac{\kappa}{1+\beta e^{-\alpha h}} \right)^2 - \sum_{h=1}^n \left(y_h - \frac{\hat{\kappa}}{1+\hat{\beta}e^{-\hat{\alpha}h}} \right)^2 \leq 3s^2 F_\alpha(3, n-3) \left(1 + \frac{5}{3} \hat{N}_\Phi \right) \quad \dots (6.13)$$

Een betrouwbaarheidsuitspraak welke een benadering is van (6.13) is

$$P \left\{ \begin{array}{l} \hat{\kappa} - \sqrt{k X'X^{11}} \leq \kappa \leq \hat{\kappa} + \sqrt{k X'X^{11}} \\ \hat{\beta} - \sqrt{k X'X^{22}} \leq \beta \leq \hat{\beta} + \sqrt{k X'X^{22}} \\ \hat{\alpha} - \sqrt{k X'X^{33}} \leq \alpha \leq \hat{\alpha} + \sqrt{k X'X^{33}} \end{array} \right\} \geq 100(1-\alpha)\% \quad \dots (6.14)$$

met

$$k = \frac{3}{n-3} \sum_{h=1}^n \left(y_h - \frac{\hat{\kappa}}{1+\hat{\beta}e^{-\hat{\alpha}h}} \right)^2 F_{\alpha}(3, n-3) \left(1 + \frac{5}{3} \hat{N}_{\frac{\alpha}{2}} \right)$$

en

$X'X^{jj}$, het j-de diagonaal element van $(X'X)^{-1}$.

De variantie rond de geschatte niet-lineaire regressie-lijn wordt als volgt berekend:

$$\text{Var}(\hat{y}_h) = u^2 \text{var}(\hat{\kappa}) + v^2 \text{var}(\hat{\beta}) + w^2 \text{var}(\hat{\alpha}) +$$

$$2uv \text{cov}(\hat{\kappa}, \hat{\beta}) + 2vw \text{cov}(\hat{\beta}, \hat{\alpha}) + 2uw \text{cov}(\hat{\kappa}, \hat{\alpha})$$

waarbij

$$u = \frac{1}{1+\hat{\beta}e^{-\hat{\alpha}h}}$$

$$v = \frac{-\hat{\kappa}e^{-\hat{\alpha}h}}{(1+\hat{\beta}e^{-\hat{\alpha}h})^2}$$

$$w = \frac{\hat{\kappa}\hat{\beta}he^{-\hat{\alpha}h}}{(1+\hat{\beta}e^{-\hat{\alpha}h})^2}$$

7 Appendix Scheffé F - projectie [†]

Lemma 1.

De lineaire variëteiten

$$M_1 = \{y \mid a'(y-c) = r(a'a)^{\frac{1}{2}}\} \quad \dots (A.1)$$

$$M_2 = \{y \mid a'(y-c) = -r(a'a)^{\frac{1}{2}}\} \quad \dots (A.2)$$

zijn evenwijdig, staan loodrecht op de vector a en raken een bol met straal r en middelpunt c .

Bewijs

1° Voor het bewijs dat M_1 en M_2 twee evenwijdige lineaire variëteiten (hypervlakken) zijn, loodrecht op de vector a , kunnen we verwijzen naar Hadley, G. 12, blz. 197, 199. Het volgt uit het feit dat a de normaal vector is, elke hypervlak loodrecht staat op zijn normaal vector, en twee hypervlakken evenwijdig zijn, als ze dezelfde normaal vector hebben.

2° Het raken van deze hypervlakken aan een bol. Hiertoe zullen we bewijzen dat zowel M_1 als M_2 draagvlakken (supporting hyperplanes) zijn d.w.z.

- a) dat de doorsnede van de lineaire variëteit met de bol niet leeg is;
- b) dat de bol volledig ligt in een gesloten half-ruimte gevormd door het hypervlak.

[†] Deze appendix is gebaseerd op MILLER, R.G. 17, blz. 48-67. We danken ook Dr. van de Kerkhof, H.P.J. voor de stimulerende discussies.

De vergelijking van een bol met straal r en middelpunt c is:

$$B = \{s \mid (s-c)'(s-c) \leq r^2\} \quad \dots (A.3)$$

waarbij $s = (s_1, \dots, s_p)$, een punt in of op de bol is.

a) Te bewijzen: $M_1 \cap B \neq \emptyset$;

$$M_2 \cap B \neq \emptyset$$

$$\text{Stel } b = c + \frac{r}{(a'a)^{\frac{1}{2}}} a$$

Nu is $b \in M_1$ immers deze vector voldoet aan de definitie van M_1 :

$$a' \left(\frac{r}{(a'a)^{\frac{1}{2}}} a \right) = r(a'a)^{\frac{1}{2}}$$

en

$$b \in B \text{ immers}$$

$$(b-c)'(b-c) = \left(\frac{r}{(a'a)^{\frac{1}{2}}} a \right)' \left(\frac{r}{(a'a)^{\frac{1}{2}}} a \right) = r^2,$$

voldoet dus aan (A.3)

dus $b \in M_1 \cap B$ en $M_1 \cap B \neq \emptyset$.

Indem voor $M_2 \cap B \neq \emptyset$

b) Te bewijzen: $a'(y-c) \geq r$ of $a'(y-c) \leq -r \forall y \in B$

Stel $s \in B$. Gebruik makend van de Cauchy-Schwarz ongelijkheid ($|a'b| \leq |a| \cdot |b|$), kunnen we schrijven

$$|a'(s-c)| \leq |a| \cdot |s-c| = r(a'a)^{\frac{1}{2}} \leq r \quad \dots (A.4)$$

We mogen dus ook schrijven:

$$-r(a'a)^{\frac{1}{2}} \leq a'(s-c) \leq r(a'a)^{\frac{1}{2}}$$

M.a.w.

de bol B ligt volledig tussen de hypervlakken M_1 en M_2

Theorema 1.

Indien $\hat{\beta} \sim N_p(\beta, \sigma^2(X'X)^{-1})$, X een $n \times p$ matrix met rang p , ($p < N$) en L is de p -dimensionale parameterruimte, dan is:

$$P\{\beta \mid |l'(\beta - \hat{\beta})| \leq \sqrt{ps^2 F_{\alpha}(p, n-p)} \\ (l'(X'X)^{-1}l)^{\frac{1}{2}}, \forall l \in L\} = 100(1-\alpha)\%$$

Bewijs

Aangezien $(X'X)^{-1}$ een positief definitie matrix is, bestaat er een niet-singuliere matrix D zodanig dat

$$D(X'X)^{-1}D' = I \quad \text{zie Goldberger, A.S. [10, blz.36]}$$

Stel nu

$$\delta = D\beta \quad \text{en} \quad \hat{\delta} = D\hat{\beta} \quad \dots (A.5)$$

dan is

$$\hat{\delta} \sim N(\delta, \sigma^2 I)$$

We weten (zie Goldberger, A.S. [10, blz. 178]) dat het $100(1-\alpha)\%$ betrouwbaarheidsgebied voor δ gedefinieerd wordt als

$$P\{\delta \mid (\delta - \hat{\delta})'(\delta - \hat{\delta}) \leq ps^2 F_{\alpha}(p, n-p)\} = 100(1-\alpha)\%$$

Dit gebied is dus een bol met middelpunt $\hat{\delta}$ en straal $\sqrt{ps^2 F_{\alpha}(p, n-p)}$.

Uit Lemma 1 (A.4) volgt nu

$$P \{ \delta \mid |a'(\delta - \hat{\delta})| \leq \sqrt{ps^2 F_{\alpha}(p, n-p)} \mid (a'a)^{\frac{1}{2}}, a \} = 100(1-\alpha)\% \quad \dots (A.6)$$

Uit (A.5) volgt nu:

$$\left. \begin{aligned} a'\delta &= a'D\beta \\ a'\hat{\delta} &= a'D\hat{\beta} \end{aligned} \right\} \Rightarrow a'(\delta - \hat{\delta}) = a'D(\beta - \hat{\beta})$$

$$a'a = a'Ia = a'D(X'X)^{-1} D'a$$

en aangezien D een niet-singuliere matrix is, kunnen we a kiezen zodat

$$a'D = 1'$$

(A.6) kan nu geschreven worden als

$$P \{ \beta \mid |1'(\beta - \hat{\beta})| \leq \sqrt{ps^2 F_{\alpha}(p, n-p)} \mid (1'(X'X)^{-1}1)^{\frac{1}{2}}, \forall 1 \in L \} = 100(1-\alpha)\%$$

Q.E.D.

Corollarium 1.

$$P \{ \beta_j \mid |\beta_j - \hat{\beta}_j| \leq \sqrt{ps^2 F_\alpha(p, n-p) (X'X)^{jj}} \} = 100(1-\alpha)\%$$

met $(X'X)^{jj}$, het j^e diagonaal element van $(X'X)^{-1}$

Dit is te bewijzen door een juiste keuze van de 1 vector: een 1 als j^e element en de andere componenten nul.

Literatuuropgave:

- [1] BARNES, J.G.P. "An algorithm for solving non-linear equations based on the secant method", The Computer Journal.
8e jrg., april 1965, nr. 1, blz. 66 - 72.
- [2] BEALE, E.M.L. "Confidence Regions in Non - Linear Estimation", Journal of the Royal Statistical Society, Series B.
22e jrg., 1960, nr. 1, blz. 41 - 88.
- [3] BROYDEN, C.G. "A class of methods for solving non-linear simultaneous equations", Mathematics of Computation.
19e jrg., 1965, blz. 577 - 593.
- [4] ———. "Quasi - Newton Methods and their Application to Function Minimisation", Mathematics of Computation.
21e jrg., juli 1967, nr. 99, blz. 368 - 381.
- [5] CROCKETT, J.B. en H. CHERNOFF. "Gradient Methods of Maximization", Pacific Journal of Mathematics.
5e jrg., 1955, blz. 33 - 50.
- [6] DAVIDON, W.C. "Variable Metric Method for Minimization", A.E.C. Research and Development Report,
ANL - 5990. november 1959., 21 blz.
- [7] DRAPER, N.R. en H. SMITH. Applied Regression Analysis.
London, John Wiley & Sons, Inc., 1966, 407 blz.
- [8] FLETCHER, R. en M.J.D. POWELL. "A rapidly convergent descent method for minimization", The Computer Journal.
6e jrg., juli 1963, blz. 163 - 168.
- [9] GALLAUT, A.R. "A Note on the Measurement of Cost/Quantity Relationships in the Aircraft Industry", Journal of the American Statistical Association.
63e jrg., december 1968, nr. 324, blz. 1247 - 1252.

- [10] GOLDBERGER, A.S. Econometric Theory. London, John Wiley & Sons, Inc., 1964, 399 blz.
- [11] GUTTMAN, I. en D.A. MEETER. "On Beale's Measures of Non - linearity", Technometrics. 7e jrg., november 1965, nr. 4, blz. 623 - 637.
- [12] HADLEY, G. Linear Algebra. London, Addison - Wesley Publish. Co., Inc., 1961, 290 blz.
- [13] HARTLEY, H.O. "The Modified Gauss - Newton Method for the Fitting of Non - Linear Regression Functions by Least Squares", Technometrics. 3e jrg., mei 1961, nr. 2, blz. 269 - 280.
- [14] HEUTS, R. "Schattingen van parameters in de gamma - verdeling en een onderzoek naar de kwaliteit van een drietal schattingsmethoden m.b.v. simulatie", Statistische Notitie nr. 7, Tilburg, december 1969, 17 blz. + bijlagen.
- [15] KENDALL, M.G. en A. STUART. The Advanced Theory of Statistics, Vol. 2, Inference and Relationship. London, Charles Griffin, 1967, 2e uitg. 690 blz.
- [16] MALINVAUD, E. Statistical Methods of Econometrics. Amsterdam, North - Holland Publish. Co., Studies in Mathematical and Managerial Economics, Vol. 6, 1966, 613 blz.
- [17] MILLER, R. G. Simultaneous statistical inference. London, Mc Graw - Hill Book Company, 1966, 272 blz.
- [18] ROSEN, E.M. "A Review of Quasi - Newton Methods in Nonlinear Equation Solving and Unconstrained Optimization", National Conference of the ACM, Proceedings of the 21st Conference. Washington. D.C., Thompson Book Co., 1966, blz. 37 - 41.

- [19] SHANNO, D.F. "Conditioning of Quasi - Newton Methods for Function Minimization", Center for Mathematical Studies in Business and Economics, University of Chicago, Report 6910 (Revised), augustus 1969, 20 blz. (voorgelegd ter publicatie aan Mathematics of Computation).
- [20] ——— en P.C. KETTLER. "Optimal Conditioning of Quasi - Newton Methods", Center for Mathematical Studies in Business and Economics, University of Chicago, Report 6937, augustus 1969, 16 blz.
- [21] SVESHNIKOV, A.A. Problems in Probability Theory, Mathematical Statistics and Theory of Random Functions. London, W.B. Saunders Company, 1968, 481 blz.

Bibliotheek K. U. Brabant



17 000 01059110 6

E.I.T. 1970